Patterns in Language, Culture, and Society:
Sub-Saharan Africa. OSU WPL 19.125-141 (1975)

The Beginnings of Ethnohistory in Western Wellegga: The Mao Problem*

M. L. Bender
Southern Illinois University, Carbondale

1. Introduction

Most of the fourteen provinces of Ethiopia extend outward from the central province of Shewa like spokes from the hub of a wheel. Wellegga is one which extends due west to the Sudan border.

The western-most part of Wellegga consists for the most part of lowlying swamps and semi-desert beyond the central Ethiopian plateau. However, this is broken by a strip of higher land paralleling the Sudan border from north to south through the major towns of Asosa, Bambeshi, Gidami, and Dembi Dolo. The higher land (about 2000 meters) gets more rain and is fertile agricultural country, while the lower (about 1000 meters) is generally arid and cannot support a high density of population.

The main ethnic groups in western Wellegga are Arabs, Berta, Galla (Oromo-speakers), Gumuz, Komo, Kwama, and "Mao". In addition there are significant numbers of "central Ethiopians" (Amhara farmers, Amharic- and Tigrinya-speaking schoolteachers, etc.). See the map on the following page.

Linguistically, these groups are native-speakers of languages of two of the four great African super-families identified by Greenberg (1966). Arabic and Amharic are Semitic and Oromo is Cushitic, all members of the Afroasiatic superfamily. Berta, Gumuz, Komo, and Kwama are members of the Nilo-Saharan superfamily. According to the Greenberg classification, all these except Berta, in addition to the language of the "Northern Mao" belong to the Koman family of Nilo-Saharan, while Berta is one of four branches of the Chari-Nile family. The "Southern Mao", a few hundred agriculturalists living in the Anfillo forest west of Dembi Dolo (see map) speak a language which belongs to the Kefa group of Omotic, one of the six Afroasiatic families. (In Greenberg's classification, Omotic was considered as "West Cushitic": for the reclassification, see Fleming (1969) and Bender (1974)). "Southern Mao" (henceforth: Anfillo) is rapidly being displaced by Oromo. Figure 1 presents family trees:

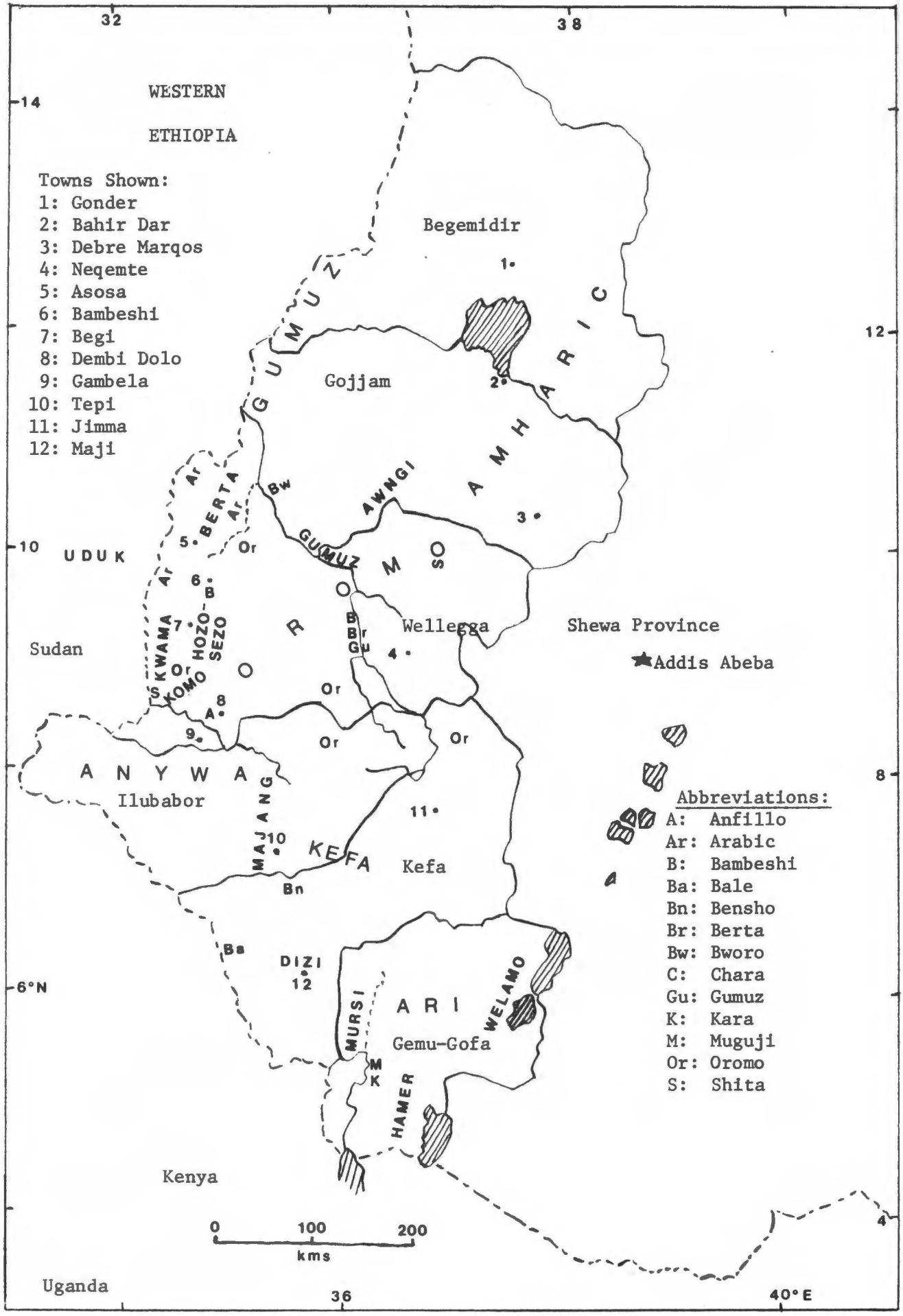
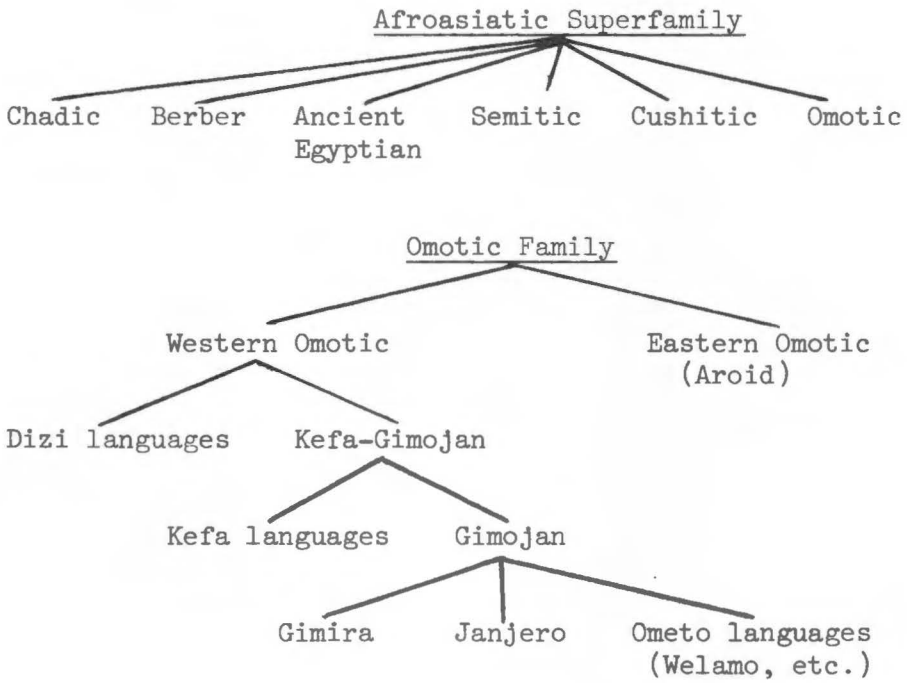

WESTERN
ETHIOPIA
Towns Shown:
1: Gonder
2: Bahir Dar
3: Debre Marqos
4: Neqemte
5: Asosa
6: Bambeshi
7: Begi
8: Dembi Dolo
9: Gambela
10: Tepi
11: Jimma
12: Maji
Begemidir
Gojjam
Wellegga
Shewa Province
Addis Abeba
Sudan
Kenya
Uganda
UDUK
ANYWA
Ilubabor
KEFA
Kefa
DIZI
ARI
Gemu-Gofa
GUMUZ
AMHARIC
AWNGI
MAJANG
MURSI
WELAMO
HAMER
SKWAMA
HOZO-SEZO
KOMO
Abbreviations:
A: Anfillo
Ar: Arabic
B: Bambeshi
Ba: Bale
Bn: Bensho
Br: Berta
Bw: Bworo
C: Chara
Gu: Gumuz
K: Kara
M: Muguji
Or: Oromo
S: Shita
0 100 200
kms
32
38
14
12
10
8
6°N
4
36
40° E

Figure 1: Simplified Family Trees



1. Nubian
2. Surma: Mursi, Majang, Muguji, etc.
3. Nera (Barya)
4. Ingessana (Tabi)
5. Nyima, Afitti
6. Temein, Teis-um-Danab
7. Merarit, Tama, Sungor
8. Dagu
9. Nilotic: Anywa, etc.
10. Ngangea, Ik

The evidence for subclassification of Nilo-Saharan languages in western Ethiopia is slight: little data was available to Greenberg on Berta or Koman languages. Much more data has become available in recent years, and the result is that some problems have been solved while others are now seen to be more difficult than at first supposed.

For the present, the classification of Berta, Komo, Kwama, and the Afroasiatic languages of the area will be taken as given. But Gumuz shows little in common with Komo-Kwama, and "Mao" still less.

Up to here, I have been writing "Mao" with quotation marks. This is because there is really no such language. In fact, "Mao" is a term used by Gallas to refer to the very dark-skinned people of the area, much as the general Ethiopian term "Shanqilla" is used. The meaning of "Mao" in the languages of some of the people concerned is simply "people".

The people who are usually referred to as "Mao" include the following: (1) Anfillo, mentioned above, not Nilo-Saharan-speakers (at least not now); (2) The Komo and Kwama of west and south of Begi (see map); (3) The Hozo and Sezo clans of the Begi area; (4) The Maramo clans of the Bambeshi area.

Of these, only in Hozo (ma^{w}) and Sezo (ma:y), is "Mao" or a variant of it, the word for "people". However, Bambeshi seems to be fairly closely related to Hozo and Sezo, and I shall henceforth refer to Hozo, Sezo, and Bambeshi as the Mao languages. Komo, Kwama, and Anfillo are clearly very different from all three of these, though Kwama shows a fairly high lexical resemblance to Bambeshi.

Further details concerning the ethnic and linguistic situations in western Wellegga are found in Bender (forthcoming) and Atieb, Triulzi, and Bender (forthcoming).

2. The Mao Problem

Greenberg's classification of Mao as Koman is based on the data available at the time, mainly Grottanelli (1940), which gives very little linguistic data. The Grottanelli data is on the languages of Anfillo and Bambeshi. When I first started collecting Mao in the Begi area, I assumed that it was the same language as that of Bambeshi. A little experience soon showed that this was not the case; in fact local informants had told us that there were eighteen "different Mao languages" in the Begi area!

Investigation reduced this to three: the eighteen actually enumerates clans, six Kawma-speaking, six speaking a language self-named ho:z wa:ndi, and six speaking se:zi wa:ni ("Hozo-mouth" and "Sezo-mouth"). The latter two can be considered as separate languages or as dialects of one language, according to choice of criteria. They are not mutually intelligible by the strictest criterion, but speakers of two varieties converse, each speaking his own variety and understanding enough of the other to communicate. Thus the situation is like that of Spanish and Portuguese, which are considered as separate languages more for political than for linguistic reasons. Note that Bambeshi Mao is not mutually intelligible with Hozo-Sezo: Bambeshi and Hozo-Sezo are two quite distinct languages.

When working on my lexicostatistic classification of all Ethiopian languages (Bender 1971) I was mystified to find that Bambeshi Mao (of which I had collected a 100-item list in 1969) showed much higher lexical resemblance to several Omotic languages than to Koman languages. For example, Bambeshi-Chara stood at 31%, Ometo languages (average) 20%, Gimira and Dizi languages 17% each, Janjero 13%, Kefa languages 12%, Kwama 11%, Kara 10% (Bender 1971:205).

At the time and for several years thereafter, I was inclined to assume that the high Bambeshi-Omotic figures were due to heavy borrowing during some period of extensive contact in the past. I hoped sometime to get around to check this by taking a more careful look, using fuller lexical sources, and carefully screening out loanwords. In the meantime Harold Fleming had already taken what I thought was a drastic step by deciding that Mao is Omotic and not Koman at all (in personal communications and in Fleming (forthcoming)).

It is the purpose of this present paper to try to settle the question: is Mao Koman, Omotic, or something else? I propose to look at both lexical and morphological evidence, taking advantage of the mass of new data now at my disposal as a result of the data-collecting by Fleming, myself, and others during the past decade. I see this effort as the first step in trying to unravel the complex skein of ethnohistory in western Wellegga: Who are the older inabitants of the area and who are the recent immigrants? How does it happen that there are scattered enclaves of Omotic speech far to the north of the main Omotic areas? Etc.

3. The Grammatical Sample

It was not a purpose of the present paper to provide ammunition for either side in the dispute over whether lexical or morphological evidence is more useful or decisive in deciding genetic relationships. My own point of view is that both are valuable and neither can be ruled out. The only real difficulty arises when the two kinds of evidence seem to be contradictory. One notorious example is that of Mbugu or Ma7a in Tanzania: lexically South Cushitic, but morphologically Bantu (see Goodman (1971) for a summary of the status of Mbugu). It was my suspicion that Hozo-Sezo would turn out to be another Mbugu type. To me this appeared to be a not wholly unhappy prospect, since it would provide another example for investigation into the fascinating realm of language hybridization. I am fairly sure that there are several of these in Ethiopia, in addition to the one reported by Habte-Mariam Marcos (1973 and 1975).

I thought that a look at Hozo-Sezo grammar as compared to that of representative Omotic and Koman languages might make the question of genetic relationship transparent. To do this I chose a sample of Omotic and Koman languages from major branches of the two families insofar as my data permits. I wanted to keep the samples chosen for lexical and grammatical data as uniform as possible, and this was a further limitation, though it mainly meant choosing the lexical sample first, since there are few cases in which adequate grammatical information is available but not adequate lexicon.

The languages chosen for grammatical comparison thus consist of Hozo, Sezo, and Bambeshi (henceforth HSB), Welamo, Kefa, Dizi, and Ari (representing the main branches of Omotic), and Komo, Kwama, and Gumuz from Koman. For control purposes and for various ulterior reasons, I adjoined a sample of Nilo-Saharan languages from branches other than Koman: Berta, because it is in contact with HSB; Anywa, Mursi, Majang, and Muguji because they are far removed from HSB. For details about these languages,

see Bender (1971, 1974, and forthcoming); in brief, Majang is included because of its deviance and probable isolation from outside influences; Mursi because it is an "orthodox" Surma language; Anywa because it is the major Ethiopian Nilotic language; Muguji because it is a possible hybrid. All serve as controls on HSB-Omotic and HSB-Koman comparisons: not much can be made of similarities between HSB and either Omotic or Koman if they are not significantly higher than those found between HSB and one or more of the control languages.

The choice of morphological items for comparison is again dictated by the availability of data. In this case, I abstracted from data collected for a comparative Nilo-Saharan study (Bender, forthcoming) those items which were found to be sufficiently documented for the present sample to make comparisons meaningful. These include: pronouns (independent subject and object, possessive, verb affixes), demonstratives, interrogatives, sex-indicators, noun plural markers, verbal nouns ("to" and "-ing" forms), causative and passive derived verb markers, tense markers (present, past, future), copulas (existence, place, "have"), and negative markers (verbal, copular, other).

Note that a fair number of these (from among pronouns, demonstratives, and interrogatives) are also included in the basic lexicon.

4. The Morphological Evidence

Hozo, Sezo, and Bambeshi (HSB) form a grouping in terms of lexical commonality, as we shall see in section 5: Hozo-Sezo is 64%, Hozo-Bambeshi is 37%, and Sezo-Bambeshi is 47% on a 50-item basic lexicon. None of the three scored more than 31% against any other language. In fact, as noted earlier, Hozo and Sezo are on the borders of mutual intelligibility. The figure of 64% for two marginally mutually intelligible languages agrees well with other results I have gotten in such cases (e.g. Afar-Saho, Kunama-Ilit, Tigre-Tigrinya (see Bender 1971).

In terms of grammatical morphemes, however, the case for HSB is much weaker, especially for the inclusion of Bambeshi (though even in lexical terms, Bambeshi is a questionable member). Because this paper must be kept brief, I am not able to present more than a sample of data here. I would be glad to provide full data to anyone writing to request it.

4.1. Pronominal Elements.

First, let us look at pronominal elements. The first step was to collect these under the headings: independent subject pronouns, object pronouns, possessives, and verb affixes. The following person distinctions were found necessary (in looking at HSB and other pronominal systems): first singular, second singular, third masculine and feminine singulars, third neuter or common singular, first plural inclusive and exclusive, second plural, third plural.

I made an attempt to extract a common element from HSB. I found this possible only in the cases of 2 sg. and 3 sg. as

hi and n respectively. This is a discouraging start: The pronouns of HSB seem to have as little in common as those of three totally unrelated languages!

If HSB pronouns do not give much grounds for commonality, how do they succeed in perhaps linking up all or part of HSB with Omotic, Koman, or other Nilo-Saharan? Wherever possible, "common Omotic", "common Koman", and "common Surma-Nilotic" forms were set up for comparison with HSB. The most convincing comparisons are:

Hozo-Sezo 3 sg. a, common Omotic 3 sg. a
Hozo 1 pl. nu, common Omotic 1 pl. (i)no
Bambeshi 1 sg. ti(ŋ), common Omotic 1 sg. (in)ta

Other HSB-Omotic similarities can be found, but they strike me as less convincing or involving single languages. I would like to caution against getting carried away by either my or other examples because of (a) the subjectivity of such comparisons, (b) the small number of possibilities inherent in pronoun systems, thus leading to a high level of chance coincidences. As an example of the latter, consider the English and Hamer pairs: 2 sg. you, ya; 1 pl. we, wo; 2 pl. you (older ye), ye. Hamer is an Aroid language.

To sum up, I consider the three stronger HSB-Omotic correspondences listed above and the many weaker ones (some very weak indeed!) as being insignificant in establishing an HSB-Omotic connection.

Is the case any stronger for HSB-Koman? Here I find the following most interesting:

Sezo 3 sg. han, common Koman h
Bambeshi 1 pl. ham or m, Koman m
Sezo 2 pl. nam or m, Bambeshi (h)aw, Koman m
Hozo 3 pl. ine, ena, Koman n?

These are more numerous but not more convincing than the HSB-Omotic ones. They are based on forms found (usually) in only one HSB language and two Koman languages, and always depend on a single consonant.

How about HSB with Surma-Nilotic? Consider:

Hozo 1 sg. na, common S-N ani
Sezo 1 sg. vb. af. a, common S-N a
Sezo 2 sg. hin, n, common S-N inu
HSB 3 com. sg. n, common S-N ene
Bambeshi 3 com. sg. iš, iša, S-N š
Bambeshi 3 pl. išk, common S-N k.

These are still more in number, but again unconvincing. The -n- of Hozo hin may be a result of assimilation to a suffix, the -k- of Bambeshi išk is poorly attested. These weaken an already weak case based on generally one HSB language and coincidences of single consonants. However, the presence of both 3 com. sg. n and 3 com. sg. š in both HSB and Surma-Nilotic is interesting.

Note that 3 sg. s is found in Welamo and z in Dizi. HSB comparisons with Berta yield no convincing pairs. In fact Berta is highly idiosyncratic compared to all the other languages. The HSB-Berta contact situation has not led to "intimate" borrowings of pronouns as in the English-Scandinavian case.

My conclusion to this section must be that the findings are only suggestive: in comparisons of HSB with Omotic, Koman, and Surma-Nilotic, it seems that the similarities cannot be placed above the level of insignificance in any case. In fact the use of Surma-Nilotic as control languages has resulted in a standoff: HSB does not look any more like Omotic or Koman than Surma-Nilotic, and it does not look much like any of them in terms of pronouns.

4.2. Other Grammatical Elements

These consist of demonstratives, ("this" and "that"; plurals turned out to be of little use), interrogatives (who?, what?, when?, where?), masculine and feminine sex-determiners (e.g. to indicate "Billy-goat" and "Nanny-goat"), verbal noun ("to" and "-ing" forms, often identical), noun plural markers, causative and passive markers in verbs (data on other derivations was too scanty), imperfect and perfect aspect markers (most of the languages in question seem to have aspect rather than tense systems: the markers were not differentiated in the comparisons, sometimes they mark completed vs. incompleted action, sometimes present-future vs. past, etc.), copulas (existence, place, and possession: often two or all three of these are the same), and negative markers (used with verbs, copulas, and others, e.g. adjectives, adverbs). Of the above, the demonstratives and interrogatives are also found among basic lexicon.

Here more common HSB elements can be set up than with pronouns. The common forms for "what?" kin- and "where?" hint- are impressive; most of the rest are Hozo-Sezo only. Comparing (common) HSB with (common) Omotic does not lead to many strong correspondences.

Similar remarks can be made for HSB vs. common Koman, S-N, and Berta. The comparisons are not enough to convince me of anything, but one interesting pattern which seems to be above the level of chance is that of frequency of Bambeshi, Majang, and Muguji in these comparisons. At present, this can be only suggestive: consider this together with the facts that Bambeshi is clearly quite distinct from Hozo-Sezo, and that Majang and Muguji are isolated and low-prestige groups among the Surma-speakers.

4.3. Conclusions

I did not try to quantify these grammatical comparisons because the data is not full enough nor in all cases of high-enough quality. Thus my conclusion will be more subjective and more open to attack than I like, but on the other hand, perhaps someone will see through the maze and demonstrate that more definite conclusions are possible on the basis of this data. Ultimately I would like to see the data base expanded and improved so that a quantitative classification in the manner of Chretien and Kroeber (1937) (for Indo-European) could be drawn up.

Given the limitations on quality and quantity of data, the problems of using such generally tiny bits of phonological substance as grammatical morphemes generally are, the limited possibilities in pronominal systems, etc., and the apparently random distribution of correspondences found, I conclude that we better go on to lexical comparisons if we hope to approach a solution of the HSB problem.

5. Lexical Comparisons

5.1. Numerals

I hadn't planned to start off this section with a discussion of numerals, but after finishing the grammatical analysis the thought suddenly occurred to me that numerals might provide a clue since they are both part of the basic lexicon and notoriously subject to borrowing (especially higher ones). Thus I compiled Table 1 comparing the first five numerals and the word "hand" (because of its similarity to "five" in many cases) for all the languages used in the lexical study. The result was so interesting and shows so clearly the problems inherent in trying to classify HSB that I have decided to present the table of numerals here.

TABLE 1
NUMERALS

	1	2	3	4	5	hand
Hozo	onna	dombɔ	siyazi	bɛts'i	k'witsi	kutsi
Sezo	ɪšilɛ	no:mbɛ	si:zɛ	bɛ:ts'ɛ	ʊwi:sɛ	kusɛ
Bambeshi	ɪški	yumbo	te:ze	mɛts'ɛ	wʊsi	kusɛ
Welamo	ista, isɔ	na7a	he:za	oyda	icɛc, ša	k^huši
Chara	issa	nanta	ke:za	oβda	u:ca	kuca
Bensho	ma:t'	nəm	kəz	o:d	ʊ:joc	kuc
Anfillo	iko	guto	ke:djo	awdo	ammɪtto (Sem.)	ke:šo
Kefa	ɪkko	gu:to	ke:ymo	awɪddo	u^wɪco	ki:šo
Dizi	k'oy	ta:gn	kàdú	kùbḿ	úx́cú	kucu
Ari	wəlɛk'a	k'astɛn	məkkən	oydi (Om.)	dɔ:nk'	ani
Berta	(mʊ:)ŋk'ú	(mó:)làŋ	(mú:)θè	(mè)némù	(mé)k'ú:šú	θàbá
Majang	o,umun	pe:y	ji:d	aŋan	tu:l	a'ri
Bale	ɔDe	raman⁻	i:yɔ	we:	turr	a:yi
Muguji	kiɛm⁻	Da	jiɛn⁻	ɛho7	cɔn	mbu:a
Shita	diše-De	šukā	tušu	hwan	muta-kwei	kwoye
Komo	De7	so, swɛ́yén	diš(én)	dwɔ́gɔ́n-(én)	bʊ́s'(én)	k'wolo
Uduk	De7	su7	kwara	doŋon	muDeD	meD
Kwama	sɛndo, sɛ:nɛ	si:ya	twasan	bi:s'ɛn	kʊm-but'	bit'
Gumuz	meta	mba:nda	koga	nzi:k^ya	ma-k^wo:sa (Berta?)	ela

The first three languages of Table 1 are HSB; following are seven Omotic languages. To the previous selection I have added Chara, from the same group as Welamo (the Ometo cluster), Bensho, giving a wider perspective on the "Gimojan" branch (see Figure 1), and Anfillo, another "Kefoid" language, but a particularly critical one since the Anfillos are the "Southern Maos", people who may once have been speakers of our HSB languages. The reason these were not included in the grammatical comparisons is that these are among the languages for which lexical data is available but not grammatical.

Looking first at the numerals "one" to "five", it doesn't take long to discover that all the HSB numerals are cognate to either Omotic or Koman numerals except for Hozo "one": onna (this seems closer to English than to any of the given Ethiopian languages!).

Thus, HSB falls together with Omotic on one, two, five. "Three" seems to be temptingly similar to "two" in Koman, especially when compared to Hozo siyazi, Kwama si:ya. This comparison is not to be ruled out, but I have accepted as more likely the direct comparison with Koman "three" through Bambeshi te:ze, Kwama twasan to include Shita and Komo also. An argument can also be made for HSB with Omotic through s, t, -z against K-Z, but I think the Koman case is stronger. I am reluctant to accept a t-k correspondence without further supporting evidence. Dizi kàdú, Uduk kwara is also suggestive: perhaps ultimately all are cognate forms (and one might then bring in Majang ji:d, Muguji jien~, Afroasiatic S-D).

"Four" presents the weakest case: HSB B-s' vs. Kwama bi:s'ɛn (the rest of Koman given in Table 3 is quite different). Certainly Omotic presents an unlikely appearance (a die-hard might want to argue for Omotic W-d in this connection, perhaps from *b-d, see Chara).

[Many important details have had to be omitted here.]

I have not attempted to analyze numerals beyond five because they shed little new light on the problem at hand.

I have looked at numerals in some detail because they indicate clearly the mixed nature of the HSB lexicon: considerable affinity with Omotic, especially Ometo (Welamo and Chara in my sample) and Bensho, but also significantly with Koman. The corresponding items do not give the overall appearance of recent loanwords in either case: practical identity rather than significant reshapings indicating some considerable passage of time since either borrowing or a period of genetic commonality. The Koman correspondences seem "deeper" in this sense than the Omotic ones, but the sample is small and subject to question.

5.2. Basic Lexicon

Passing now to general lexicon, a few preliminary explanations are in order. I had intended to do lexical comparisons at three levels, using 100-item lists at each level: basic lexicon, less basic lexicon, cultural lexicon, hoping thereby to sort out correspondences due to genetic relationship from borrowings.

(My original prejudice was to prove that HSB is a Koman cluster with heavy borrowing from Omotic. As I went ahead with this work, and as I listened to Harold Fleming's repeated warnings, I had less and less confidence in this assumption).

Primarily because of the magnitude of the job, and secondarily because of incomplete lists, I reduced the size of the undertaking to 50 basic items, 50 less basic, and 50 cultural. Later I decided to postpone the second 50 perhaps indefinitely, because I began to wonder whether they would add anything significant to the results.

The "first 50" were chosen from my latest basic 100 list (overlapping about 85% with the Swadesh-Rea list; see Lehmann (1962:112-13)) by choosing 27 of the 55 nouns, 10 of the 19 verbs, 8 of the 16 adjectives, and 5 of the 10 other words on that list. I tried to keep the original "mix" by choosing proportional parts of semantic categories (such as body parts, natural phenomena, pronouns, etc.) and I otherwise preferred those items which were present on all lists or which had given me least trouble in eliciting (in terms of synonymy, vagueness, etc.).

The "second 50" were chosen from the second 100 of my revised 200-item list in the same way. The 50 cultural items were chosen from my cultural items questionnaire in order to keep the list as complete as possible for as many languages as possible (with HSB being given preference).

To make a long story short, I then went through several steps to prepare a table of proposed plausible cognates in the format shown by the sample (Figure 2).

Figure 2: Sample Cognations. Item No. 24 "heart"

1	5	10
Ho nibba } (Amh.?)	C tefa	Br bɪ̌sí
An yɪbbo }		
2	6	11
Se šini	Bn ma:y	S kʉ́bkānā
Ba sini	Ma Ba:y	Gu kuBitsa
Mu še:n		
3	7	12
Bm - ɛŋɛ	Kf mu:llo	Ko du:
4	8	13
W wəzɛna	D cuonu	Kw ši:ši
	9	
	Ari bu:da	
Loans	Missing	Questions
Ho, An < Amh	U (Compound)	2, 4, 8, 14; 6, 7

Here 13 plausible cognate sets are shown for basic item no. 24, several having only one member each, i.e. consisting of isolated

elements. Below the list are given probable loan words and sources, languages for which the items are missing in the available word-lists (or purposely omitted becuase they are synonymous with other items on the list, or compounds) and questions as to possible further combining or cognate sets, e.g. for item 24, it is suggested that cognate sets 2, 4, 8, and 14 might be combined.

Loans were sought for in the main languages in contact with HSB and the other languages: Arabic, Amharic, and Oromo. This task has been carried out rather hastily and incompletely and needs to be redone.

The results of the comparisons are summarized in Table 2, which is a cross-table of percentages of presumed cognates between pairs of languages for the "first 50" items. A similar cross-table of percentages on the 50 cultural items is omitted because of lack of space.

TABLE 2
PERCENTAGES OF PLAUSIBLE COGNATES: "FIRST FIFTY" BASIC LEXICON

Se	Bm	W	C	Bn	An	Kf	D	Ari	Br	Ma	Ba	Mu	S	Ko	U	Kw	Gu	
64	37	26	26	23	24	13	15	7	2	0	0	5	7	6	5	5	7	Ho
	47	28	27	25	26	15	15	7	2	0	2	9	9	10	11	9	4	Se
		31	31	31	26	19	20	11	4	4	0	4	6	4	4	12	4	Bm
			55	41	49	32	24	13	4	2	2	9	6	6	2	8	0	W
				44	38	38	14	15	0	4	2	9	2	4	0	4	2	C
					35	27	28	9	0	4	2	9	4	6	2	6	0	Bn
						53	16	7	0	4	0	2	2	6	4	7	0	An
							21	14	0	4	0	11	0	4	0	7	2	Kf
								11	2	2	6	11	4	10	10	4	8	D
									4	13	11	9	0	4	0	2	2	Ari
										6	11	2	4	4	4	4	8	Br
											27	13	6	4	0	0	4	Ma
												23	6	10	9	4	2	Ba
													2	9	11	4	0	Mu
														41	35	25	11	S
															69	29	8	Ko
																40	10	U
																	12	Kw
																		Gu

Names of languages:

Hozo
Sezo
Bambeshi
Welamo
Chara
Bensho
Anfillo
Kefa
Dizi
Ari
Berta
Majang
Bale
Muguji
Shita
Komo
Uduk
Kwama
Gumuz

Addenda:

Hozo-Bworo 28
Sezo-Bworo 26
Bambeshi-Bworo 22
Bworo-Kwama 6
Hozo-Mocha 17
Sezo-Mocha 24
Bambeshi-Mocha 29
Hozo-Janjero 17
Sezo-Janjero 21
Bambeshi-Janjero 28

Table 2 is blocked off to show significant groupings. As mentioned earlier, HSB forms a group of some sort, with an I.L. (internal low: lowest between-language percent in the group) of 37 and an E.H. (external high: highest percentage of any of HSB against an outside language) of 31 (Bambeshi vs. Welamo, Chara, or Bensho). Likewise Majang-Bale-Muguji (part of the Surma family) form a weak group with I.L. of 13 and E.H. of 11, except for Majang-Ari, which is 13. Shita-Komo-Uduk-Kwama (part of the Koman family) has I.L. of 25 and E.H. of 12 (Kwama-Gumuz).

Gumuz is classified by Greenberg as a Koman language, and indeed Gumuz shows its highest lexical relationships with Koman. But if we extend the Koman group listed above to include Gumuz, the I.L. slips to 8 (Komo-Gumuz) while external figures of 8 through 12 (E.H.: Kwama-Bambeshi) are obtained. Because of the small size of my lexical base (50 or fewer items), a change in only two items could make a difference of as much as 8% in the cognation percent. We have to admit that the case for including Gumuz with Koman is very weak.

The same problem emerges in trying to include Ari in an Omotic group. In fact, no fully satisfactory Omotic group larger than Welamo-Chara or Anfillo-Kefa emerges without extending it to include HSB.

In brief, we have to extend the Omotic grouping to include all of HSB, Welamo, Chara, Bensho, Anfillo, Kefa, and Dizi. Even this is not quite satisfactory because it has an I.L. of 13 (Hozo-Kefa) and an E.H. of 15 (Chara-Ari). But the inclusion of Ari makes matters worse: as with Gumuz in Koman, Ari is certainly very deviant within Omotic if it is a member. Looking at grammatical data, one will also find that Gumuz and Ari show fewer grammatical points in common with their supposed larger groupings than do the other members. This is borne out by the findings with a 100-item list (Bender (1971:190) for Ari; see Table 12 for Gumuz-Sese-Disoha within Koman (Bender 1971:176)). Sharpening of these results might be accomplished by extending the lexical base (this is a good argument for doing the "second 50" computations). Another direction to go for improvement is a more thorough search for loan words.

The important result for present purposes is that lexically HSB must be considered part of Omotic: in terms of basic lexicon, excluding loan words as best as possible, HSB scores significantly higher against Omotic (range 15 to 31 excluding Ari) than against Koman (range 4 to 12) or the control languages (0 to 9), Furthermore, as Fleming (forthcoming and personal communication) points out, an assumption that borrowings could account for the HSB-Omotic relationship would imply that a language can have up to 31% of its basic lexicon as borrowings. This would make HSB a contender for all-time champion among borrowers! In fact Swedesh (1964:632) provides a mathematical demonstration that borrowings cannot exceed 15% in common over a long period of time.

It is worthwhile looking at a few rectangular blocks in the Table also. These show infra-group relationships as against intra-group (which are triangular blocks). First, consider HSB against

"traditional" Omotic, i.e. the block HSB vs. Welamo, Chara, Bensho, Anfillo, Kefa, Dizi, Ari. This clearly divides again: HSB vs. Ari is significantly lower than the rest (this is part of the problem of including Ari in Omotic, though Anfillo-Ari also falls to 7%). The block HSB vs. Kefa and Dizi is next, much higher than HSB vs. Ari but much lower than HSB vs. Welamo, Chara, Bensho, Anfillo. It is interesting that Anfillo is in the highest group, whereas Kefa is in the middle. Could this be related to the fact that the Anfillo are the "Southern Mao" (in other words, that they have a special ethnic relationship to HSB?). But note that Anfillo also scores significantly higher against Welamo, Chara, Bensho than Kefa does. Then look at the horizontal slice Bambeshi against traditional Omotic as compared to Hozo and Sezo: Bambeshi scores consistently higher.

Another rectangle to consider is HSB against Koman: but here one has to include Muguji, in order to say HSB acores consistently higher against the block than against the control languages (Berta, Majang, Bale). Since we are in the area of small percentages here, the question of significance comes up strongly: the Muguji result may be a fluke. One could look here, as elsewhere in problematical cases, at the particular correspondences to see whether they might be knocked down. But this could be dangerously circular: since all "cognations" were supposedly done by the use of uniform criteria, we cannot be justified in juggling figures to make things work out right, except where outright errors are found.

5.3. Cultural Lexicon

The problem of significance is severe here, because the number of comparisons is generally so small (range 10 to 38, average probably about 20). This means that a difference of only one presumed cognate more or less could result in a percentage difference of 10, or even more. Thus we should be really cautious about interpreting these results.

Comparing the results with those obtained for basic lexicon above, it is first necessary to point out that Bambeshi, Anfillo, Shita, Komo, Uduk, and Gumuz are not available for comparison. It is unfortunate that Bambeshi is unavailable and also only one Koman language is available, and that one is the one in closest contact with HSB. However, more data may be forthcoming in the near future and the results could be extended (if nothing else, Uduk could be laboriously filled in from the Uduk-English dictionary of Beam and Cridland).

Most of the corresponding cognate percents appear to be little changed or constant (e.g. Chara-Dizi 14 in both cases, Kefa-Berta 0 in both) while others show increases or decreases. Most of these can be discounted when the small size of sample is carefully studied. Space is too restricted to present any treatment of cultural lexicon here. In general, the results are quite parallel to that of the first 50 items.

Contrary to my original hypothesis that HSB would show few correspondences with Omotic in basic vocabulary, but many identifiable as loans in cultural vocabulary, the HSB-Omotic

correspondences appear to be pervasive and not identifiably loans. In fact the identifiable loans are mostly from Amharic, Arabic, and Oromo. Plausible loans from Omotic are rare.

5.4. A Note on Bworo

After doing most of the work on this paper, I realized that I had omitted a potentially important Omotic language. The northernmost Omotic language appears to be the dialect cluster known as "Shinasha" (Bender 1971) or "Bworo" (Fleming forthcoming), belonging to the Kefa group. This consists of a number of enclaves on the northern side of the Blue Nile in Gojjam Province, across from Berta, Gumuz, and Oromo-speaking areas of Wellegga. The sample of Bworo I obtained from an informant in Metekkel, Gojjam, in 1969, showed 46% common basic lexicon with Mocha, 42% with Kefa, 48% with Anfillo (Bender 1971:175, Table 8) on a 100-item list.

Since Bworo is geographically proximate to the "Mao" area, it might be interesting to see how it compares with other Kefa languages in comparison to HSB. I have added these figures to Table 2 (based on the present 50-item list). As can be seen, Bworo resembles Anfillo in its degree of relationship to both HSB and to Kwama. Bworo is also closer to Anfillo than to any other language, though well below, e.g., English-German.

5.5. A Note on Ganzo

The Ganza or Ganzo of the Sudan border area south of Jebel Bange (a Kwama area) reported by Bryan (1945:192) can hardly be other than the Hozo-Sezo. Reasons: first, the location; second, the fact that Bryan found at least 17 of 62 (27%) and as many as 30 of 62 (49%) lexical items in common with "Western and Omo Sidama."

6. Conclusions

Figures like 26% or 31% for basic lexicon shared by languages not in immediate contact are very hard to "put down" as proof of genetic relationship. In fact, as noted earlier, Swedesh provides a convincing argument that the upper limit for shared basic vocabulary between genetically non-related languages is 15%.

Search for loanwords did not reduce the plausible cognations between HSB and Omotic languages: most identifiable loans are from "culture-bearing" languages of the area (Arabic, Amharic, Oromo) and affect both HSB and Omotic.

HSB seems to be a member of the Gimojan branch of Omotic: highest correspondences are with Ometo and Bensho (see figure 1). What about the third member of Gimojan, Janjero?

A quick check gives the figures addended to Table 2. Janjero is seen to be intermediate between Kefa and the rest of Gimojan, but tending toward Gimojan, as far as its lexical relationship to HSB goes. However, Mocha is even closer to Gimojan, as seen by the addended figures in Table 2. Taking the problem of significance into account, I think it is safe to say that HSB "goes with Gimojan," and is somewhat more removed from Kefa

languages, with Bworo and Anfillo being special cases because of geographical proximity.

What about the comparisons, running 4-12, between HSB and Koman languages? These are as high as HSB-Ari, but are right in the range of probable diffusional influence.

In basic lexicon terms, the case is quite clear: HSB are Omotic languages; similarities to Koman are diffusional. This is not so for grammar. Here the evidence indicates a slight preponderance toward Omotic, but no very convincing case as to HSB belonging to either Omotic or Koman can be made on the basis of our present state of knowledge of HSB grammar.

Finally, how did the HSB languages get where they are? I think the answer to this will have to be seen in the light of the larger problem of outlying areas of Omotic speech as represented by the Bworo cluster, HSB, and Anfillo. The distribution suggests that of a relic area: the remains of a once wide-spread Omotic family in western Ethiopia, split up by Nilo-Saharan incursions from the west, and later Oromo invasions from the east. A large-scale comparative study of grammar and lexicon, especially cultural lexicon, is called for.

Footnote

*This paper is a direct outcome of work done under National Science Foundation grant no. GS-37483: Nilo-Saharan Linguistics, awarded to me through Southern Illinois University, Carbondale.

Travel funds were provided by Southern Illinois University for me to attend the Sixth Annual African Linguistics Conference at Ohio State University to present this paper. I would also like to acknowledge the aid of the following in supplying data and useful comments: Edward Allen, Atieb A. Dafallah, Harold Fleming, Morris Goodman, Wendy James, Jean Lydall, Dennis Tully.

References

Atieb Ahmed Dafallah, M. L. Bender, and Alessandro Triulzi. Forthcoming. Berta, in Bender (Ed.) Forthcoming.

Bender, M. L. 1971. The languages of Ethiopia: A new lexico-statistic classification and some problems of diffusion. *Anthropological Linguistics* 13.5:165-288.

_____. 1974. *Omotic: A New Afroasiatic Language Family.* Southern Illinois Museum Series No. 3, Carbondale.

_____. Forthcoming. *The Ethiopian Nilo-Saharans.* Artistic Printing Press. Addis Ababa, Ethiopia.

_____, (ed.). Forthcoming. *The Non-Semitic Languages of Ethiopia.*

_____, J. D. Bowen, R. L. Cooper, and C. A. Ferguson, (eds.). 1974. *Language in Ethiopia.* Oxford University Press.

Beam, Mary S. and A. E. Cridland. 1970. *Uduk-English Dictionary.* Linguistic Monograph Series No. 4, University of Khartoum.

Bryan, M. A. 1945. A linguistic No-Man's Land (The Ethio-Sudan border). *Africa* 15.4:188-205.

Chretien, C. D. and A. L. Kroeber. 1937. Quantitative classification of the Indo-European languages. *Language* 13:83-104.

Fleming, Harold C. 1969. Classification of West Cushitic within Hamito-Semitic. in Daniel McCall, et. al., *Eastern African History*, 3-27. Prager: New York.

_____. Forthcoming. Omotic overview. in M. L. Bender (ed.), *The Non-Semitic Languages of Ethiopia.*

Goodman, Morris. 1971. The strange case of Mbugu (Tanzania). in Dell Hymes (ed.), *Pidginization and Creolization of Language*, 243-254. Cambridge University Press.

Greenberg, J. H. 1966. *The Languages of Africa.* Indiana University Press.

Grottanelli, V. I. 1940. *I Mao, (Missione Ethnografica nel Uollega Occidentale* Vol. 1), Reale Academia d'Italia, Rome.

Habte-Mariam Marcos. 1973. Pidgin Italian? *Journal of the Language Association of Eastern Africa.* 3.1.

_____. 1975. Italian. in Bender, et. al. (Eds.), 170-180.

Lehmann, W. P. 1962. *Historical Linguistics: An Introduction.* Holt, Rinehart, and Winston.

Murdock, G. P. 1959. *Africa: Its Peoples and Their Culture History.* McGraw-Hill.

Swadesh, Morris. 1964. Diffusional cumulation and archaic residue as historical explanations. in Dell Hymes (ed.), *Language in Culture and Society*, 624-637. Harper and Row.

Tucker, A. N. and M. A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of Northeastern Africa.* Oxford University Press for I.A.I., London.